Cybernetics and Systems: An International Journal, 32 : 11–51, 2001

0196-9722/01 $12.00 + .00

LIKE ME? – MEASURES OF CORRESPONDENCE AND IMITATION

CHRYSTOPHER L. NEHANIV
KERSTIN DAUTENHAHN

Adaptive Systems Research Group, Faculty of Engineering and Information Sciences, University of Hertfordshire, Hatfield, United Kingdom

Imitation is a powerful mechanism for efficient learning of novel behaviors that both supports and takes advantage of sociality. A fundamental problem for imitation is to create an appropriate (partial) mapping between the body of the system being imitated and the imitator. By considering for each of these two systems an associated automaton (respectively, transformation semigroup) structure, attempts at such mapping can be considered (partial) *relational homomorphisms*. This article shows how mathematical techniques can be applied to characterize how far a behavior is from a successful imitation and how to evaluate attempts at imitation arising from a particular *correspondence* between the imitator and model.

For the imitator and the imitated, affordances in the agent-environment structural coupling are likely to be different, all the more so in the case of dissimilar embodiment. We argue that the use of what is afforded to the imitator to attain corresponding effects or, as in dance, sequences of effects, is necessary and sufficient for successful imitation. However, the judged degree of success or failure of an attempted behavioral match depends on some externally imposed or—in the case of autonomous agents—internally determined criteria on effects of the attempted imitative behavior (including effects attained successively as well as final effects). These criteria correspond to *metrics*—measures of difference—which can guide the evaluation of a correspondence, the learning of a correspondence, or learning how to apply one. Metrics on states and sequences of action events in the system-environment coupling allow judgment of similarity for 'observer-dependent' purposes. This allows one to formally define successful

Address correspondence to C. L. Nehaniv, Department of Computer Science, University of Hertfordshire, College Lane, Hatfield, Herts AL10 9AB, UK. E-mail: CL.nehaniv@herts.ac.uk

imitation with respect to such criteria. The resulting measures can be used to compare various candidate mappings (e.g., body plan or perception-action correspondences). Additionally, this may be applied in the automated construction and learning of mappings to be used in imitation for artificial, hardware, and software systems.

A starting point for an inquiry into imitation is the question: *Are you like me?* That is, is there a sufficient degree of correspondence between imitator and model for imitation to be possible.

We are going to use the term ‘‘imitation’’ in a general way to include observational learning, copying, matching, ‘‘true imitation,’’ goal emulation, etc. All of these will be seen to be examples in which a correspondence problem has been or is being solved. Novelty of the behavior is not required, but a degree of novelty will necessarily be present if any form of social *learning* has taken place.

Imitation, matching of behaviors, and observational learning among agents engaged in social interaction can provide an efficient and powerful means for acquiring new behaviors for robots and software artifacts just as they do for many animals. These mechanisms not only take advantage of social interaction but can also support learning and culture in a community of agents. Automation of the matching of behaviors, looking for correspondences between embodied systems and their ways of interacting with their surrounding environment, and specific mechanisms for attempting imitation offer inroads to the application of principles from the worlds of biology and developmental psychology to robotics. To clarify the problems and key points in constructing systems that try to imitate, we introduce a mathematical framework in which to address these issues, with a view not only toward engineering applications in robots and other types of artificial agents, but also as a larger unifying framework that serves as a guide in which imitation in animals and artifacts can be systematically addressed. Wherever mechanisms of imitation and related behavior are to be described or designed, a unified systematic viewpoint is possible despite differences among the various scientific disciplines.

First, we shall distinguish some modes of imitation. We discuss how the purpose of an attempted imitation relates to its degree of success. We address how observer-attributed or autonomously generated purpose relates to abstraction about action and perceptions, and how this can shape the meaning of ‘success’ for imitation. From this basis, we go on to consider automata (and associated algebraic structures) that describe the imitators and systems being imitated. This leads to the notion of 'relational homomorphism' between systems, which formally captures the idea of attempted correspondence between systems for imitation. To evaluate an attempted correspondence, an additional notion is necessary: the effect of sequences of actions and/or the effects in terms of state changes in the environment in the course of the imitator's behavior must also be 'close' to those effected by the model's behavior. Measuring similarity (for a given purpose)—using (pseudo)metrics—of the various states and the various action-event sequences in system-environment couplings allows the formal characterization of imitative behavior successful for that purpose.

The plan of this article is to identify the key issues in correspondence between dissimilar bodies, affordances of the body-environment coupling, and the role of observer criteria in behavior matching judged to be successful imitation. While relegating mathematical details to Part II, in Part I we motivate, overview, and discuss the formal framework in regard to imitation and related aspects of behavior-matching and observational learning in animals and artifacts. This framework is intended to be general in that it supports the precise comparative study of imitation (and matching behavior in general) in various artificial and natural systems, as well as providing a mathematical language for the engineering application of imitation in the realms of robotics, machine learning, and software. Relational homomorphisms will be used to capture the intuitive notion of *correspondence*, even between agents with dissimilar embodiments and different sensory and effector apparatus with different degrees of freedom, etc. Metrics will quantify the degree of difference or similarity in sequences of action events and resulting states performed for some (in general, not intentional, and possibly externally attributed) purpose or goal of the behavior.

PART I:
IMITATION

Action-Level, Program-Level, and Effect-Level Imitation

Whether one considers imitation between organic, embodied, virtual, synthetic, software, or more abstract systems, it is possible to distinguish types of imitation varying with respect to the manner in which events (actions, but also sensations) of an imitator "correspond" to those of the imitated system. The question of what *correspondence* is, is central for imitation. Which actions of an imitator do or do not correspond to those of a model is largely observer-dependent. Indeed, whether or not a particular behavior was a successful attempt at imitation must be answered with reference to correspondence criteria, in terms of whether certain [sequences of] effects on the agent-environment coupling were attained.

Under the simplifying assumption that the bodies (and couplings to the environment) of both systems are of the same structure, the problem of mapping actions is a trivial one once the (extremely difficult) perceptual problem of identifying homologous body parts and actions is solved: ‘‘Replaying’’ the same actions on the body of the imitator, an action of one system is mapped to the same action on the other.

Byrne and Russon (1997) distinguish *action*- and *program-level* imitation. In action-level imitation, the imitator carries out exactly the ‘‘same’’ actions as the imitated system. The actions are described by a detailed, linear specification of sequential acts. In program-level imitation, the imitator carries out an identical ‘‘program’’ (overall structure, sequence of steps). Such a program is conceived of as a structure of hierarchical subroutines. A step in a program may itself invoke the execution of a sequence of subroutines. Thus, the *structure of knowledge* used in producing a matching behavior has a program-like—as opposed to association-like, or string-like—form. Byrne (1999) has emphasized the possibility of the extraction of such programs by animals via string parsing and detection of statistical regularities. Whiten (1999) has carried out experiments with children and chimpanzees to study whether and in what way hierarchical procedures are learned socially. Work on programming by example (e.g., Cypher, 1993; Lieberman, 2000) and behavioral cloning (e.g., Sammut et al., 1992) has been geared at enabling software agents to acquire program-level knowledge by observing human performance of such skills as lay out text and pictures or flying a simulated aircraft. Various researchers have classified different forms of ‘‘imitation’’ and social learning, according to whether actions, results, or goals are matched (see especially Call and Carpenter in press for a discussion of these three sources of information in social learning).

Throughout this article, we allow but shall never be require that bodies of the imitator and its model (the system being imitated) have the same structure. Once bodies are no longer of identical type, what could be meant by the "same" actions or program is no longer unambiguous. (Together with factors of noise, individual differences, and differences in circumstances, even with bodies of the same type in similar circumstances, the notions of "same" actions and "same" program are certainly not unproblematic.)

As an approach to the correspondence problem for dissimilar bodies, we suggest consideration of the problem at a more abstract level without making assumptions about the structure of knowledge acquired in social learning. Imitation has an essential aspect of *form-filling* in which the *effects* on the environment and the body of the self are to be mimicked. Actions are never performed in a vacuum, conceptually isolated from their embodied presence and effects in the world (although some research work on imitation has endeavored to treat it as such). These effects may be described in terms of actions, states, or goals, and combinations of these. Achieving corresponding effects by acting is possible by bringing about changes of state in the environment, in the positioning, movements, and orientation of own's own body, and also by performing similar actions with one's effectors. We shall be primarily concerned with what it means for two behaviors by two different actors to correspond. Secondarily, we shall mention in passing some possible forms that the structure of knowledge that makes use of such a correspondence may take.

Many different programs can lead to the same effects. However, copied actions or programs carried out in inappropriate contexts fail to achieve desired effects (e.g., strokes in painting far from the wall aren't achieving any painting!). This notion of effect-level imitation is similar to that recognized as "functional imitation" by Demiris and Hayes (1997).

Achieving an effect, as discussed below, depends on how an agent uses *affordances* (see Gibson (1977), Norman (1988), and below) in the structural coupling between itself and its environment. Intuitively, affordances are the perceived or actual properties of how something may be used by the agent. This coupling, and hence these affordances, and hence the actions carried out in taking advantage of them, may be quite dissimilar for the two agents even during successful imitation. Learning tool- and action-affordances can be the result of observational learning or imitation.

From this perspective of achieving corresponding effects, it is natural to consider *following* behavior, as it has been used in various robotic experiments (Hayes & Demiris, 1994; Dautenhahn, 1994; Billard & Dautenhahn, 1997, 1998; Billard & Hayes, 1997; Gaussier et al., 1998) to be a simple form of imitation. The *path* or at least the *successive locations* of one agent are approximately matched by the follower. The follower need not have a body very similar to that of the leader for one to make sense of the notion of 'occupying roughly same series of locations in the same order.' The follower may have radically different sensory input (e.g., infrared versus touch sensors) and different effectors (e.g., wheels versus limbs). An example of learning by imitating using following is illustrated in Figure 1.

Successful following means that the locations are matched successively and thus the traces of the paths of the two agents match. For some purposes, such as reaching a desired destination (e.g., a charging station whose location is known to the leader but not the follower), the imitation can be considered successful if the behavior of the follower takes it to the goal.

The similarity of the paths and whether the changes in velocities in the course of the following approximately match or not could be further criteria in considering the closeness of the imitation. More extremely, in dance, position, posture, velocity, and acceleration changes should all be matched by the follower imitating the dance instructor. In matching behavior for many purposes, such as, for example, in dance, successful imitation may depend on how tightly a sequence of subgoals in terms of all these (and other) variables is matched.

At effect level, imitation or matching behavior requires a match of the effects produced by the imitator and model. Effects may be *external* in relation to the body, influencing the environment and

---

Fig 1. (*opposite*) The photos show two robots that were used in the experiments described by Billard and Dautenhahn (1997), where a learner robot while following and imitating the movements of a teacher robot learns a vocabulary to label its perception of inclination, e.g., to distinguish between moving up and down a hill and moving on plain ground. The top photo shows the teacher (left) and the learner robot (right) in the initial position. Following is implemented with infrared sensors. The robots are not identical; they have different shapes and sensory-motor characteristics. The experiment assumes that the teacher robot "knows" how to interpret the world, i.e., it is emitting different signals (bit-strings, by radio-link communication) for moving on a plane and moving on a hill. This is shown on the bottom photo. The learner robot has to learn the teacher's interpretations of "words" on the basis of its own dissimilar sensory perceptions. Learning proceeds via the creation of associations, using a recurrent neural network architecture DRAMA developed by Aude Billard (1999).

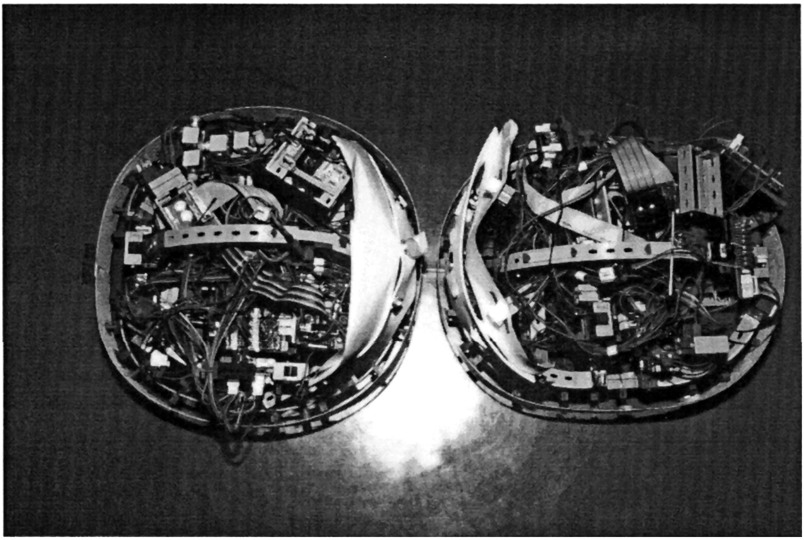

Figure 1 (top)

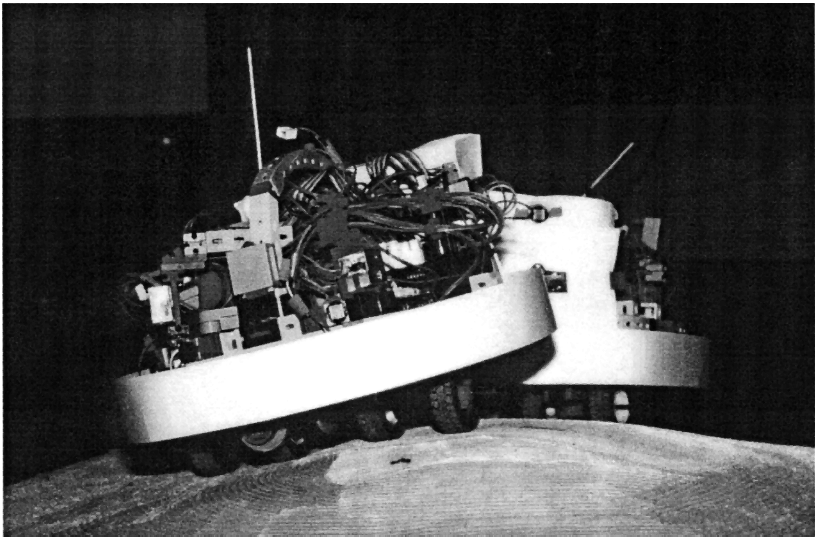

Figure 1 (bottom)

objects in it (e.g., opening a fruit, producing a song, kicking a ball, painting a wall, computing the sum of a list of numbers), or be *internal* (e.g., facial expressions, tongue protrusion, body postures, joint angles), or both (e.g., performing a swan dive, dancing, camouflaging oneself as a prey animal or innocuous environmental component).

DISSIMILAR BODIES

Finding appropriate mappings between bodies of individuals of the same kind (species) can be relatively straightforward, e.g., when two robots share the same set of basic movements like forward, backward, left, right, then actions can be mapped accordingly. Mappings between dissimilar bodies depend far more on the observer point of view, determining the level of comparison between movements, e.g., selecting the [sequences of] subgoals of the movements (influenced and contextualized by their effects on the environment) which should be mapped.

Humans and robots obviously have dissimilar bodies. Differences always exist between humans (e.g., between adults and children, and due to individual variation in bodies and behavior), between humans and other animals, and between robots. Figure 2, for example, describes a scenario where robots with different sensors and knowledge use following to help a 'blind' robot get to a charging station in a study of social interaction. Even commercially available robots of identical type have differences in sensory-motor characteristics. Distinct individuals are never the same. In fact, difference in embodiment is more the rule than an exception and must be recognized in any satisfactory theory of imitation.

Following an automobile with manual transmission in one with automatic transmission, the precise actions of the drivers need not (and cannot) match exactly: it is, for example, impossible to 'let out the clutch' in the car with automatic transmission. But, nevertheless, it is possible to match approximately the path and even the changes in velocity during following. The precise actions and behavioral programs are different for the two drivers but the form to be filled, the achievement of a certain matching trajectory, is still possible despite the dissimilarity between the two systems and the actions that their controllers (the drivers) must exert on them to

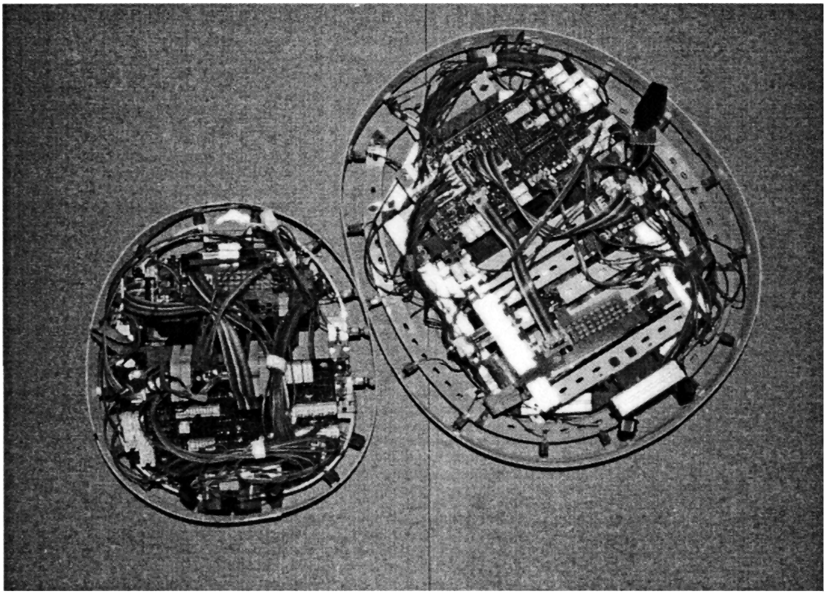

Figure 2. Dautenhahn (1997a) describes an experiment with a helping scenario between heterogeneous robots: a robot equipped with light sensors serves as a "seeing-eye" robot which can guide a second "blind" robot to a charging station (visible by a bright light source). The "blind" robot does not have light sensors but can use its contact sensors in order to follow the "seeing-eye" robot. The photo shows the follower (right) and the smaller guide robot (left). Size, weight, speed, and sensor equipment of the robots are different. A ring of 16 contact sensors is attached to the white material around the robots. The image shows the "blind" robot making contact with the "guide."

achieve corresponding behavior. One sees immediately from this example that consideration of imitation is both possible and useful for differently embodied systems. The mapping of actions is no longer trivial, as someone who learns to drive a manual transmission after years of driving with automatic transmission will attest, but such a mapping is certainly achievable.

To give another example: If a humanoid robot is trying to interact and imitate a human, the exact morphology and fine movements of its hands are unlikely to match those of the human partner (e.g., different degrees of freedom for the robot and human arms). However, if the desired behavior is turn-taking, then for this (social) interaction the goal is the successful, dynamic coordination of the movements rather than the precise copying of the human's movements.[1] This imitation (from the observer perspective) does not even require the representation of imitation behavior. As experiments with the Cog project (Brooks & Stein, 1994; Brooks, 1996; Brooks et al., 1999) show, turn-taking between a human and a humanoid robot can be an emergent property of the interaction dynamics (Brooks et al., 1999; Breazeal, 1999, 2000; Breazeal & Scassellati, in press)

As the similarity between bodies and their manner of acting on the environment decreases, the problem of mapping becomes, in general, more and more difficult and finally degenerates into impossibility. A human's body might enact motions corresponding to those of a horse, but less successfully those of an insect or worm (due to excess or surplus of corresponding legs!). A bird's actions in flying might be mapped from the pectoral muscles and wings into corresponding motions on a subset of muscles and fins on a fish, but this does not necessarily enable the fish to fly. However, the hovering or forward motion of a hummingbird can be imitated by something as dissimilar as a helicopter, by mapping wing orientation to the tilt of the rotor. Then in both systems, for hovering, lift cancels weight (Figure 3, top pair), while a degree of tilt in the wings or rotor creates a component of forward thrust as in the lower pair of Figure 3 (cf. McNeill Alexander, 1992; pp. 119–125). The correspondences between angle of the rotor and angle of the bird's wings and between changes in these two angles are the essential part of a mathematical correspondence called a 'relational homomorphism' between the helicopter and the hummingbird, which serves as a recipe for generating matching behaviors for purposes of hovering or of mov-

---

[1] As discussed elsewhere (Dautenhahn, 1999) with respect to robot-human interaction, dynamic coordination of movements need not even require a stable mapping of corresponding actions and body parts (i.e., correspondences might change during the course of interaction), as long as the *temporal coordination* is preserved, namely, the temporal structure of movements of two agents are matching sufficiently. To give an example: A group of dancers in a disco can each perform very different movements but as a group they can still give the impression that their dance is coordinated with each other (and mediated via the music rhythm, although temporal coordination does not necessarily need external synchronizing stimuli). Similarly, successful imitation need not necessarily involve a fixed mapping correspondence, but may involve switching between several mappings or a relational correspondence compatible with several mappings. Moreover, temporal matching of speed and duration of actions during successful imitation may or may not be required, depending on observer criteria. Finally, a correspondence may be learned, constructed, or modified in the course of interaction.

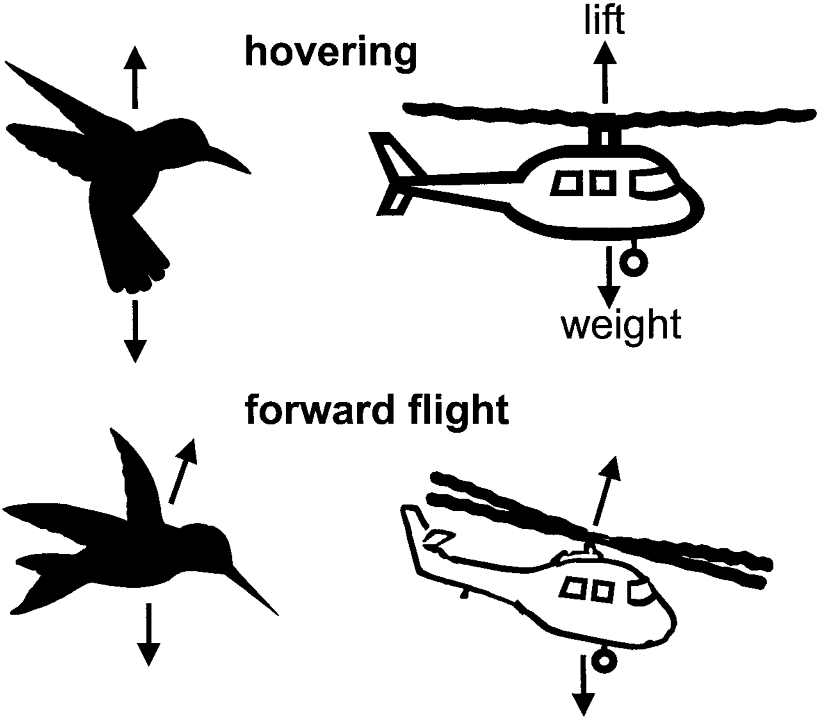

Figure 3. Top pair: Stable hovering in hummingbirds and helicopters employs the same principle: exactly cancel the effect of weight using lift. This technique also occurs among other (necessarily small) birds and insects that can hover. Lower pair: With rotor at an angle and vertical component of lift cancelling weight, a horizontal force component moves the helicopter forward. This same principle of tilting through an angle is used in the forward (and backward) flight of birds and insects that can hover in flight, despite the fact that the mechanisms to generate motion differ radically, i.e., rotation of rotor versus flapping of wings.

ing backwards and forward while maintaining altitude. Moreover, this relational homomorphism is 'good' since observer criteria—and associated metrics—show that corresponding changes in each system lead to corresponding effects in the other: under this correspondence, changes in altitude and direction are judged to be similar in quality and degree, if not scale.

If bodies and their environmental couplings are too dissimilar, then successful imitation can be impossible. For instance, presumably no amount of creativity could lead to finding successful correspondences for the actions involved in a human's brushing her teeth and the movements of a radically different system such as a bicycle used in getting from point A to point B.

Effects: Imitation for Some Purpose

An aspect of effect-level imitation is concerned not necessarily with the specific actions the imitator performs, but that the result should *in some particular sense* be the same. Imagine a child with a thin brush. It is required to imitate an adult human who is painting a wall with a wider brush. Corresponding motions by the imitator will not reliably cover the wall with paint. Contact of the brush with the paint and the wall will be required of the imitating system, just as for the adult painter. The difference in height of the agents implies the imitator will have to compensate for this difference, perhaps by lifting its painting arm through a great angle, by extending it disproportionately to how the adult human is extending her arm, or possibly it will necessitate that the imitator first lifts itself up (e.g., by climbing a ladder) to reach all places that need painting. The difference in brush width requires more strokes to cover the same area. Action- or program-level scaling of action would not cover the desired area if it were not concerned with effects outside the body on the environment. One sees that goal *effects* can be in *state* (covering in the wall)—in this way the new state is a desired *result*, and can also be in *action* (e.g., mimicking arm movements), or of mixed nature (e.g., learning to hold a brush correctly). In Part II, this will be formally captured by applying separate metrics on states and on action-event sequences, while also permitting metrics that depend on both in simple or complex ways.

The apprentice painter with its smaller embodiment would not be judged to have successfully imitated its teacher unless the purpose were also fulfilled. The form to fill here is the effect of covering the desired area of a wall (and no part of an adjacent floor or window) with paint. The (observer-attributed) *purpose* of the imitation is to achieve a similar effect on the environment, in this case the wall and surrounding objects. Matching motion for motion would look like imitation, successful by another criterion, but *not* successful according to the usual one for painting. Meanwhile, there are a huge number of action sequences and programs that can yield the same effect, e.g., holding the brush in the left or right hand, painting with one leg over the arm holding the brush, etc.

Abstraction of Action

The preceding discussion of purpose and effect in imitation leads us to the following insight. Motion-for-motion or action-for-action matching—even if novel—are not necessarily sufficient for the success of attempted imitation. In particular, such types of mapping are unlikely to be sufficient in the case of differently embodied systems. When one considers imitation for some purpose or sequence of subgoals, filling the form or achieving the effect desired for each subgoal is what is required and is sufficient for success.[2] Abstraction away from particular microactions to classes of actions that achieve the same subgoal leads to an appropriate notion of what to imitate.

> *An action or sequence of actions is a successful component of imitation of a particular action if it achieves the same subgoal as that action. An entire sequence of actions is successful if it successively achieves each of a sequence of abstracted subgoals.*

Notice that success of imitation, thus defined, depends on identification of subgoals in the behavior to be imitated. Thus, an implicit characteristic of imitation is mapping for some purpose (or sequence or hierarchy of purposes). At the same time, the imitator should avoid unnecessary effects on the environment (a following robot that takes shortcuts by smashing through walls or knocking over furniture would hardly be considered successful for practical purposes). In the painting example, a successful mapping is one that uses the body-environment coupling to get the same abstracted effects of covering the wall (and nothing else) with paint.

The judgment of success or failure is observer-dependent, since the goals may be externally imposed by an observer not within the system itself. We emphasize that in using the shorthand notions of goals or des-

---

[2] One would also like to stipulate the *nonattainment* of side effects: e.g., wild flailing of arms or getting paint on the glass of an adjacent window would be considered unsuccessful components of an attempted imitation for painting behaviors. They achieve other significant effects than those required for the purpose of painting.

ired effects, we have not abandoned the nonteleological stance of science; nevertheless, we do wish to highlight the necessity of considering (if there is one) the observer (who may perhaps be the agent itself) making the judgments on the degree of success or failure of imitative actions for a system coupled with its environment. The role of the observer is generally clear to workers in biology and ethology, who are observing animal behavior, while in the engineering disciplines the dominant view is to adopt an "objective" external viewpoint, in which observer-dependent issues are dismissed. How observer-attributed (or internally generated) goals give rise to metrics is discussed in the section on the correspondence problem and attempted imitation.

Perceptual Correspondence

The 'helping scenario' in Figure 2 exploits the *following strategy* (Dautenhahn, 1997a): the following robot is able to follow the guide to the charging station, but is not able to imitate this behavior in the absence of the guide, due to a lack of correspondence in the perception of environmental features which are necessary to find a path to the charging station. This is different from experiments that employed the following strategy for learning by imitation experiments (e.g., Hayes & Demiris, 1994; Billard & Dautenhahn, 1997, 1998; Billard, 1999). In these cases the guide and follower robot could have different sensory-motor abilities; however, the follower could perceive environmental features that were necessary in order to learn and later execute the behavior it learned from the guide. This is different from the 'helping scenario,' where the guide and the following robot have insufficient correspondences in their perception of the world. The follower has no way to detect the charging station. Thus, unless the follower would be equipped with additional sensors, the follower cannot imitate the behavior of *going to the charging station*: there is no way to build a correspondence from the follower's experience of states and events to that of the guide which can perceive a light source on top of the charging station. We suggest that, for calling a behavior imitation, a *correspondence of perception* (including extero- and proprioception) between model and follower needs to exist. Without such a correspondence, one agent might still follow or copy the behavior of another agent. But unless it has perceptions (in whatever modalities) that correlate to those that the model needs to perform the behavior, there can be insufficient correspondence and, hence, no imitation for purposes of that particular behavior. Independent of the degree of practice, repetition of the behavior, or the social relationship between the two agents, the agent could never perform the behavior alone. An external observer not aware of the architecture of the follower might wrongly believe that it could go to the charging station without the help of the guide, or, if not, that it might be able to learn based on its experience following the guide how to find and go to the charging station. Although correspondences of form, structure, degrees of freedom of the body (structural congruence), dynamics of the actions performed (temporal coordination and temporal contingency (Dautenhahn & Werry, 2000)), and correspondences of the behavioral repertoire are important aspects and might be used to classify and distinguish degrees of correspondence in imitative behavior, imitation seems to require that model and imitator share *perceptions of a shared context* (e.g., a spatial context in the case of robots following each other and learning to associate features of the environment with the imitated action). A shared context need not be spatial, e.g., an agent spatially separated from the model but able to monitor the model's behavior in its environment can still be a successful imitator. Shared context can be fixed ('designed' by nature, mimicry, or engineered), accidental (agents meeting each other at particular places), opportunistic (kin and friends, agents in social groups more generally are more likely to share a common context), or actively established (a student seeking a teacher or a teacher wanting to instruct a student). Once two agents have a corresponding perception of a shared context (possibly using different types of sensors and sensor modalities and perceiving different features of the environment), a constructed correspondence can potentially result in imitative action, depending on the degree to which the imitator can identify and execute corresponding mappings of state and events. Correspondence of perception is, therefore, a necessary (but not sufficient) condition for imitation to take place if such perception is necessary for carrying out the behavior to be imitated.

This issue of perceptual correspondence and that of action correspondence are two aspects of event correspondence treated uniformly via the event-sequence correspondence component of relational homomorphism in Part II of this article.

AFFORDANCES FROM A SITUATION

Following Gibson (1977) and Norman (1988) we refer to the perceived and actual properties of a situation (body-state, environment, stimuli, available tools, and so on), especially those pertaining to how they can be used to do something as *affordances*. Affordances depend not only on tools and their design, but also on their embeddedness in the environment and on the particular bodily structure and configuration of an agent who might use them. The looped ear of a coffee mug affords grasping for humans and many other primates with grasping hands, but not for a bird, snake, or dolphin. Imitation for an effect involves the perception of, or at least the employment of, affordances for achieving some goal, filling some form, or attaining some functionality. How a system can imitate another depends on what its embodiment and environment affords it. Often, the notion of affordance cannot be cleanly factored into system and environment, but may have an emergent character depending on coupled interaction.

Effect-level imitation, especially for differently embodied agents, relies on the use of affordances in the imitator's own body, tools, and environment to attain effects corresponding to those attained by the agent being imitated. Note that these considerations are applicable not only to humans, animals, and robots, but also make sense in the context of software and virtual agents.

An affordance makes it possible for an action that would otherwise have other results to have results useful to the agent. For example, a lever makes it possible for the act of pushing to move a large and heavy obstacle that a man pushing directly would not be able to budge. A bottle or cup provides an affordance that can make grasping and lifting into actions that provide liquid nourishment. Not only tools, but also properties of the rest of the environment can provide affordances. A cave, tree, or tall grass have affordances that make normal locomotion and resting into behavior that can yield a beneficial effect: by entering them an agent may be sheltered from predators. Flat terrain, as opposed to hilly, can afford energy savings in locomotion.

Part 1 of Figure 4 shows the trajectory of a model robot (a) and four trajectories which imitators perform (after observing or following the model). In this case it is difficult to decide from an observer point of view which the best imitator is. In Part 2 of Figure 4, the environment in the form of hills is shown in addition to the trajectories (cf. the hilly landscape scenario with mobile robots in Dautenhahn (1995)). It now

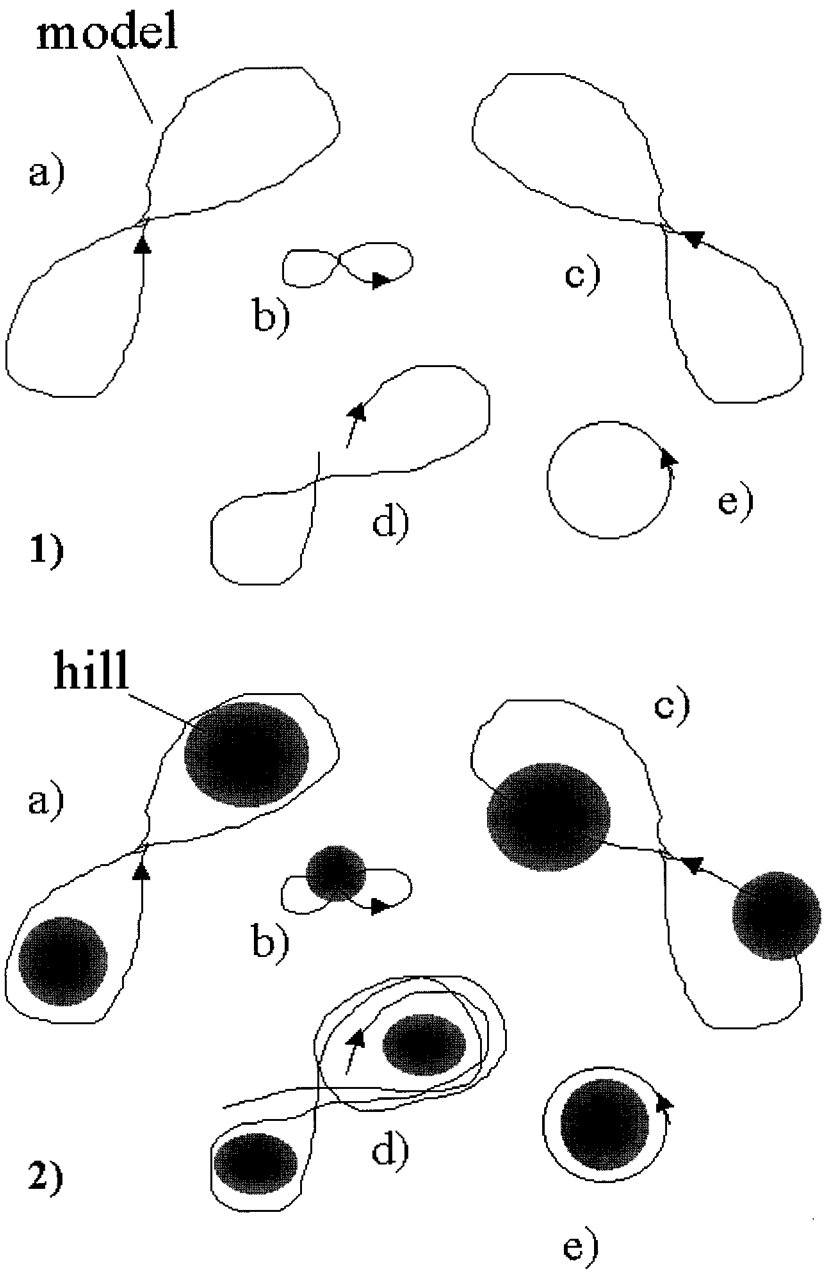

Figure 4. 1) Trajectories of a model robot (a) and four imitators (b)–(e). 2) The model and imitators in a hilly landscape.

becomes clearer that (d) and (e) might be considered good imitators, since they copy the strategy of the model, namely, ‘moving in valleys’ although the shape of the trajectory of (c) is more similar to (a) than (e) is to (a). This example shows how important the environment can be for evaluating imitative movements. If the robots are able to detect hills, e.g., by means of inclination sensors, then such an evaluation can also be done by the robots themselves. Moving in valleys reflects in this case an energy-saving strategy and is as such beneficial for the survival of the robot.

As should be obvious by now, the role of the observer is central in judging success of imitation: an observer interested in minimizing energy consumption while navigating the terrain would prefer (d) or (e). In contrast, an observer interested in matching the shapes of trajectories would judge (b) and (c) as more successful. The case where the observer is ‘internal,’ i.e., the imitator itself, is a special case and facilitates autonomy in evaluating attempts at imitation.

THE CORRESPONDENCE PROBLEM AND ATTEMPTED IMITATION: OVERVIEW OF RIGOROUS FOUNDATIONS

We discuss how one can go from the subjective notion of *observer-attributed goals* to a well-defined notion of *metrics*. The source of goals and subgoals depends on what the observer judges to be meaningful for the imitator. ‘‘Meaningful’’ here refers to usefulness in attaining goals—whether for homeostasis, survival, feeding, reproduction, learning, or other purposes (cf. Nehaniv, 1999b). Essentially, a metric will (numerically) judge two actions with similar effects (for the purpose in question) to be ‘close.’ The metric need not necessarily be completely known to the imitator but may, for instance, be implicit in the interaction with the environment (e.g., via reinforcement feedback, verbal commands of a teacher, observed reward or feeding, etc.). Closeness may be evaluated by observing the effects of actions and how well they fulfill a given purpose.

It is possible to give an algebraic formulation of the correspondence between (possibly differently embodied) systems and its relation to the matching of behaviors in attempts to imitate for some observer-attributed purpose. By carrying out this formalization, one is forced to explicitly state what is meant by success in imitation, to clearly identify aspects of body (or system), states (of system and environment), events (manipulations, actions, and sensory events), and sequences of subgoals to be used in evaluating the success of attempted imitation. This leads to welcome clarity in discussions of imitative behavior and identification of where assumptions or definitions may differ. Moreover, it also becomes possible to assess the degree and type of success in particular attempts at imitation. This, in turn, allows for the quantitative comparison of different imitative strategies and candidate solutions to the correspondence problem. Mathematical details of the approach are given in Part II, but the intuitions are outlined here.

The imitator and essential aspects of its environment are described as an automaton (see the section on automata below) whose state may be changed based on sensory stimuli and by taking actions. Similarly, the model is described in this way. The correspondence problem is to determine a *partial* correspondence between states and events for the imitator and those of the model. Such a correspondence is a mathematical construction called a 'partial relational homomorphism' (Eilenberg, 1976; Nehaniv, 1996). It is defined by the property that when the imitator and model are in corresponding states, then corresponding events (stimuli or actions) for the imitator and model must result in corresponding states for each of them.

As was seen above, how similar states (of system and environment) are for a particular purpose (subgoal) depends on observer criteria. If the model exhibits a behavior that achieves a sequence of subgoals, and the imitator exhibits its own behavior that achieves an equivalent sequence of subgoals, then the imitator is successful in its attempted imitation. The degree of success can be measured based on metrics (depending on the purpose—i.e., the sequence of subgoals to be attained) that quantify the differences between the states achieved and the [observer-attributed] subgoals. For example (as shown in Part II), by summing these differences, one has a measure of the degree of deviation of the imitator's behavior from a successful imitative behavior. (Metrics evaluating the degree of the match of actions—rather than or in addition to the matching of states—in the two systems may also be used, and again such metrics arise from observer-dependent criteria.) As indicated previously, "purpose" may but need not depend on intentionality in the imitator, model, or an external observer—formally, any of the metrics of various types just described may serve in evaluating the closeness of a matched behavior.

A solution to the correspondence problem, which results in matching of subgoals during behavior, will thus result in a recipe for successful imitations of the model by the imitator. Note that successful imitation solves a partial correspondence problem with respect to some observer (or metric), and its solution leads to a *generalizable* ability to imitate successfully (at least for the particular purposes of the matched subgoals). Nevertheless, a correspondence that works for imitation for some purpose might not work for imitation for another purpose, or may fail for the original purpose should unexpected factors—such as states on which the correspondence is not defined—contravene.

The granularity of matching of subgoals may be high or low. High granularity means that many subgoals in the model's behavior must be matched, while in an extreme case of low granularity only the final state of the behavior is important. This case, in which there is a global goal rather than a sequence of subgoals that is to be attained, is referred to as "goal emulation," where the particular method of attaining the goal is not copied (Tomasello, 1990). In high granular matching in actions but not subgoal states, we have something like blind action-level imitation, "copying" or "mimicry" in which the purpose of the behavior is "mere" action-matching rather than manipulation of the environment or the attainment of any particular resulting states. Program-level imitation implies the acquisition or learning of a program or plan of action which could make use of a solution to the correspondence problem when carrying out the program. The analysis here focuses attention on the correspondence in state, action, and goal aspects that comprise components of the behavioral program rather than the structural properties of such programs.

The issues of learning the possible effects of actions and the learning of an association of rewards with stimuli and actions of others all bear relevance to intent or purpose of behavior. The matching of behaviors of an 'imitator' with those of a 'model' may be evaluated with metrics arising from such motivations. But it is important to distinguish which agent, animal, or observer is the source of motivation or attributed motivation giving rise to the metrics. Often, an observer may feel that there is an intention in a machine or animal to match actions, effect state changes in its body or environment, or obtain a certain goal. Attributed intent may coincide, overlap, or differ wildly from the 'imitator' agent's intent. For instance, categories of observation and behavior that are of interest here include "stimulus enhancement" or "local enhancement," the drawing of attention by actions of others to particular stimuli or locations (Zentall, 1996, 2001). Following such enhancement, an animal attends to a stimulus or location and may discover how to perform a task through individual learning without regard to the particular method used by a model to solve the task. For example, in birds, the skill of opening milk bottles by pecking through aluminum foil caps to obtain cream spread quickly among great tits in the British Isles (Hinde & Fisher, 1951). This spread of the technique has been explained as being due to individual learning following local enhancement rather than imitation of the particular actions of other birds. (See Zentall (1996, 2001) for reviews.) Moreover, associations between objects manipulated by others and rewards they are then seen to obtain may lead to the learning of tool affordances and environment affordances. Tomasello and Call (1997, p. 308) refer to this acquisition of affordance knowledge as "emulation learning" in which "an individual is not just attracted to a location but actually observes and understands a change of state in the world produced by the manipulations of another, which may be its only way of learning that such a change of state is possible." Many disagreements concerning whether some matching of behavior is "true imitation" revolve around issues of purpose (goals and subgoals).

IMITATION AND LEARNING

Imitation is a powerful mechanism for efficient learning of novel behaviors that both supports and takes advantage of sociality (Dautenhahn, 1994). Two key aspects are learning by imitation and trying to imitate.

Learning by Imitation

Learning by imitation refers to the use of imitation to learn some behavior. The learned behavior may possibly be different but will be related to the behavior imitated. In the work of Hayes and Demiris (1994) on learning by imitation, robot following is engineered into the set-up for two robots of different body types (a small teacher and larger imitator). The following behavior is a successful imitation (in the formal sense described above). This imitation is used to facilitate learning:

the follower perceives the stimuli from straight hallways, left and right turning corners while following a teacher programmed to negotiate a maze. The behavior to be learned is the ability to navigate the maze alone. In the imitator, Hayes and Demiris (1994) use the mechanism of associating the environmental stimuli perceived while following with the observed behavior of the following *self*. Because the imitator is following something which makes appropriate corresponding actions given the stimuli, it learns to associate those stimuli with appropriate actions, that is, it learns to negotiate the maze.[3]

Trying to Imitate

Trying to imitate refers to solving the correspondence problem between one's own body and actions, and the actions perceived to be effected by another (Dautenhahn, 1994). Mitchell (1987) makes the point that solving the problem of imitating in different cases requires different levels of awareness and proposes a scale of five such levels of increasing sophistication. We do not address awareness here (but also see the related discussion of empathy by Dautenhahn, 1997), rather we have concentrated has been on characterizing formally what a solution to the correspondence problem for imitation is. Namely, it is a *relational homomorphism* yielding corresponding behavior that is successful, in the formal sense described above and in Part II, for a purpose or sequence of purposes. For an action-level imitation, trying to imitate means determining which specific linear sequence of event actions (if any) one's own body can carry out corresponding to those of an observed one.[4] For program-level imitation, it means determining a corresponding program and subroutines to those in a program in the teacher that is perceived to guide the behavior of the teacher. [The particular form of the program and whether or not such a program actually

---

[3] The imitator learns by replacing its "look-up table" (transition function) determining action based on current perception, which results in its following the teacher, by another transition function that results in the imitator negotiating the maze. Analysis reveals that the difference between the two transition functions is just the absence of the teacher from the stimuli in the second one. This suggests that, as a first-order approximation, learning by imitation could, in general, be characterized by an analogous replacement of transition functions.

[4] Note that this includes open problems for the perception of actions, possibly—but not necessarily—involving a model, representation, or "understanding" of the body of the other.

exists in the teacher is irrelevant.] For effect-level imitation, trying to imitate means discovering affordances in one's own body-environmental coupling that facilitate the attainment of a sequence of subgoals (effects) that correspond to those attained by the other. The notion of matching for particular purposes (subgoals) is implicit in any evaluation of behavior as imitation.

The Structure of Socially Acquired Knowledge

The structure of knowledge that assesses a correspondence and makes use of it in a particular instance of matching behavior may be of various kinds. It may be a program synthesized or extracted by detecting statistical regularities (hierarchically organized programs extracted, e.g., by string parsing (Byrne, 1999; Whiten, 1999), programming by example (Cypher, 1993; Lieberman, 2000) behavioral cloning (Sammut et al., 1992)). It may be a collection of weights and confidence factors in a temporal Hebbian neural network (Billard, 1999), or be hard-wired and inflexible as in the following component used in experiments on learning by imitation. It might simply take the form of (a probabilistic) look-up table associating strings of state and action-events to those of a model, among many possibilities. Subroutines or behaviors may be built up from 'chunking' and in terms of lower-level correspondences that involve more basic units (primitive actions, sensory vectors). Sequences or hierarchies of subgoals can serve to structure matching behavior.

The classical definition of imitation "learning to do an act from seeing it done" (Thorndike, 1898) involves novelty (see also, e.g., Visalberghi & Fragaszy, in press). Copying a behavior already in one's repertoire—colloquially also referred to as imitation—requires observing a behavior and acting according to an already available correspondence. This does not require any new information, other than the information from the observation. One example is following behavior. If novelty—and hence learning—is involved then there is either (1) the construction or refinement of a correspondence, or otherwise (2) the acquisition of knowledge (e.g., a behavioral program) that enables the learner to use a correspondence that the learner already has. Point (1) is concerned with mapping perception to action. Point (2) is concerned with structure beyond this concerned with how to apply a correspondence or partial correspondence. This may take the form of a set of learned weights in an artificial neural network, a behavioral program or a production system of rules, or neural physiological control connected to an animal or agent's needs, drives, and goals. No intentionality need be involved in constructing a correspondence (1) or acquiring the knowledge of how to use a correspondence (2). These are two components an agent requires when trying to imitate.

CONCLUSION

We have distinguished effect-level imitation as an abstraction that may encompass various definitions of imitation. The notion of the *success or failure of imitation from an observer perspective* depends on the attainment of a sequence of effects (or subgoals) in the body-environmental coupling. This permits the identification of various states in the system-environment and various actions/events impinging on the system into equivalence classes which are similar or equivalent (as judged by the possibly internal observer) for some purpose. Similarity is measured by observer-dependent metrics. The concept of relational homomorphism was used to formalize algebraically the notion of an attempted solution to the correspondence problem between the imitator and model. The concept of correspondence of subgoals was captured formally by the closeness of states and, respectively, events with respect to equivalence for some purpose. This led to an algebraic definition of successful imitation and quantification of degree of failure in attempts at such behavior. Evaluation of a behavior as a candidate attempt at imitation has thus been mathematically formalized in terms of automata corresponding to the state-event structures of the imitator and model with respect to metrics arising from observer-attributed correspondence for some purpose.

We distinguished learning by imitation, where a successful imitation has already been set up, from trying to imitate, where a solution to the correspondence problem for successful imitation is sought. Using any metric on the similarity of states, one has an induced measure of the 'error' for corresponding behavior that is an attempt at imitation. This error could be minimized (via evolutionary or computational learning algorithms, for example) to optimize imitative attempts and evaluate quality in experiments on imitation. To take a simple example: using the discrete metric for purpose G (see Part II), one may just count the mismatches in subgoals attained during an attempted imitation. More elaborated metrics selected for particular applications or problem areas can be employed just as well. For example, for evaluating attempts at imitating the moving-in-valleys strategy (see Figure 4), a few of the possible metrics include measuring differences in position, differences in angle and velocity, differences in deviation from direction of minimal inclination or energy consumption, metrics depending on the fundamental group of the trajectory traced so far, metrics constructed from any of these metrics as weighted sums, as time integrals of metrics (over continuous time), metrics that vary with time, combinations of metrics whose relative weights vary with attention, vary with stage (subgoals achieved) or context of the behavior, and so on.

The notion of affordances in the body-environment coupling facilitating attainment of corresponding effects by systems with dissimilar bodies pointed the way to defining and measuring successful effect-level imitation and may provide guidance in the design and operation of systems that try to imitate. It is our hope that this work can provide a foundational language for the mathematically rigorous treatment, analysis, study, and synthesis of imitation in embodied and virtual agent systems, e.g., robots and software agents. Such a framework might also provide a useful language for the study of imitation in biology and developmental psychology (see Heyes and Galef for a comprehensive overview (1996)).

PART II:

This part of the article gives the formal details of the framework for treating correspondence and evaluation of attempted matching behaviors (including imitation). Here we present the algebraic and mathematical formulation of the treatment of correspondence between possibly dissimilarly embodied systems and quantify the success of attempts at imitation, taking observer-attributed purposes into account using metrics, as we outlined in the overview of rigorous foundations in Part I.

AUTOMATA

We treat the imitator and model as automata systems. Representing the states of a system as X and action-events (sensor values, primitive actions, internal/external clocks, user-commands, feedback, etc.) as Σ, we assume a well-defined *next state* transition function α: $X \times \Sigma \rightarrow X$, which given the current state and action-event returns the next state of the system. The Cartesian product $X \times \Sigma$ of X and Σ is the set of all pairs (x,a) with x in X and a in Σ. Thus, for each state-event pair (x,a), the function α can be viewed as a look-up table which tells the new state $\alpha(x,a)$ given that current state was x and current input a. Note that the inputs a may be quite complex in structure, comprising sensor and action/event input; for instance, a may consist of a vector of sensor values including possibly feedback and proprioreceptive sensor values for internal variables. These internal variables could also encode commands to actuators, historical and autobiographical information (Nehaniv & Dautenhahn, 1998b, 1998a), and other aspects of state.

Remarks: (1) Such a representation is deterministic, and the assumption of deterministic state-transition is justified for many real-world systems, including robot controllers (Jones & Flynn, 1993; Nehaniv & Dautenhahn, 1998b, 1998a). But this does not mean that noise could not affect the input, only that once the input, possibly perturbed by noise, and the current state are known, then they determine together the next state. (2) Alternatively, instead of using traditional automata, one could employ Crutchfield's concept of *ε-automata* in which *casual states* are equivalence classes over system states having the same probability distributions over possible futures (given an observational history) and *transition events* are new observations (1994). Use of such automata together with this framework would allow straightforward extensions of this approach to possibly nondeterministic imitators and models acting in stochastic environments. Moreover, on-the-fly construction of such automata could help facilitate on-board autonomous construction of behavioral goals and imitative strategies in robots and software agents.

Algebras of Events

Finite sequences of events in Σ comprise the elements of a set Σ^*, called the *free semigroup with identity* on Σ. The *identity*, 1, is the finite sequence of events with zero events; it has no duration and no effect on any state. The next-state function has a natural extension $\alpha : X \times \Sigma^* \to X$, defined recursively by $\alpha(x, aw) = \alpha(\alpha(x, a), w)$, where $x \in X$, $a \in \Sigma$, $w \in \Sigma^*$, and where $\alpha(x, 1) = x$ for $1 \in \Sigma^*$. For simplicity of notation, we write $x \cdot w$ as an abbreviation for $\alpha(x, w)$, and $x \cdot w_1 w_2$ instead of $\alpha(\alpha(x, w_1), w_2)$, etc.

Sometimes different sequences w and w' give rise to the same state transitions,

$$\alpha(x, w) = \alpha(x, w') \quad \text{for all } x \in X,$$

and it is then useful to regard them as the same. We write $[w] = [w']$ in this case. The sequences w and w' are thus equivalent from the point of view of achieving the same effect on states, although they may differ in length and other properties. Thus $[w]$, the equivalence class of w, is the set of all sequences of inputs which result in the same effect on states as does w. For instance, equivalence class $[I]$ is the set of all event sequences that have no effect whatsoever on any system state. If the set of states is finite, then necessarily since there are only finitely many possible mappings from X to X, we have that the set $S = \{[w] : w \in \Sigma^*\}$ is itself finite. Moreover, it is not difficult to verify that if $[w_1] = [w'_1]$ and $[w_2] = [w'_2]$, then always $[w_1 w_2] = [w'_1 w'_2]$. Thus, concatenation of input strings respects the equivalence relation which identifies two strings inducing the same state transitions on X. It follows that

$$[w_1]([w_2][w_3]) = [w_1]([w_2 w_3]) = [w_1 w_2 w_3] = ([w_1 w_2])[w_3]$$

for all event-sequences $w_1, w_2, w_3 \in \Sigma^*$. Thus, concatenation induces an *associative* operation on S:

$$a(bc) = (ab)c,$$

for all $a, b, c \in S$. This operation makes S a *semigroup*. Associativity is a just a grammatical property of sequences of events in time. By definition of S, we have that a well-defined transition function

$$x \cdot s = x \cdot w,$$

where $x \in X$ is any state and w is any sequence of action-events in Σ in the equivalence class $s \in S$. The structure (X, S) arising from the system (X, Σ) is called the *transformation semigroup* of the system. For brevity, we shall refer to elements of S, which are equivalence classes of sequences of action-events, as *events*.

State information. Since we are interested in effects on the environment, it will be necessary to include in this state set, not only aspects of the agent's state information, but also the state description of all relevant features of the environment. The state description should carry enough information to determine whether or not a state satisfies criteria on a goal or subgoal for whatever purpose (mathematized via metrics, see below).

CORRESPONDENCE: RELATIONAL HOMOMORPHISMS

A sequence of elementary action-events for system (X, Σ) given by a word $w \in \Sigma^*$ "successfully matches" a sequence $z \in \Delta^*$ in another system (Y, Δ), if w achieves the "same effects" as z. In this section this intuition shall be made more precise.

Relational Homomorphisms

Let (X, S) be the transformation semigroup of the system (X,Σ) and let (Y, Δ) be another system with transformation semigroup (Y, D). Recall that members of S are equivalence classes of finite sequences of action-events from Σ and that members of the same equivalence class induce the same mapping on the states. Similarly, D is the set of equivalence classes of sequences from the inputs Δ.

The states and action-events of the two systems may be as similar or different as one can imagine. Different numbers of degrees of freedom (DOFs) may be possible for the two systems; their states and action-events may include vectors of real, ordered, and/or discrete values, be represented by abstract data types, or physical configurations, and so on.

A *relational homomorphism* $\varphi : (X, S) \to (Y, D)$ is a pair of relations $\varphi_1 \subset X \times Y$ and $\varphi_2 \subset S \times D$ such that

$$(x, y) \in \varphi_1 \quad \text{and} \quad (s, z) \in \varphi_2 \Rightarrow (x \cdot s, y \cdot z) \in \varphi_1.$$

"If state x is related to the state y and an action sequence s is related to an action sequence z, then the state obtained from x by performing s is related to the state obtained from y by performing z."

The relational homomorphism is called *algebraic* if also

$$(s, t) \in \varphi_2 \quad \text{and} \quad (s', t') \in \varphi_2 \Rightarrow (ss', tt') \in \varphi_2.$$

"If event sequences s and t correspond for the two systems and so do event sequences s' and t', the combined sequences ss' and tt' must also correspond."[5] Relational homomorphisms are used elsewhere in mathematics and computer science, but are interpreted here as recipes for generating *attempted imitation* since they give *formal correspondences* between systems.[6]

A *correspondence* or *mapping* to model (Y, Δ) from imitator (X, Σ) is a relational homomorphism: $\varphi : (X, S) \rightarrow (Y, D)$. This formalizes the notion of correspondence. We shall *not* assume the mapping is algebraic nor that it is fully defined.

The mapping is the inverse of a function if it specifies exactly one state (respectively, event) for each state (respectively, event) of the model. The mapping is an inverse partial function—also called a *covering*—if at most one state is specified as corresponding in the imitator to any state for each state of the model, and the analogous condition holds for events. These cases are particularly useful since then the candidate imitator's matching behavior is completely constrained by what the model does. It is a *one-to-one correspondence* if both φ_1 and φ_2 are.

Exhibited Behavior and Corresponding Behavior

Let an *exhibited behavior* of system (Y, Δ), consisting of elementary action-events $t_1, \ldots, t_m$, which take it from state y to y' $(y, y' \in Y$, $t_i \in \Delta$, $1 \leq i \leq m$, $m \succ\ 0)$, be given.

For correspondence φ, an *attempted imitation* or *matched behavior* by system (X, Σ) of this behavior at granularity k, is a factorization in D of $t_1 \ldots t_m$:

$$t'_1 \cdots t'_k = t_1 \ldots t_m$$

---

[5] We might also require, as in some usual definitions of relational homomorphism (cf. Eilenberg (1976) and Nehaniv (1996)) that the relations are *fully defined*, that is, that each elementary action $s \in \Sigma$ be related to at least one element of Δ, and, similarly, each state of the first system must be related to at least one state of the second. (These conditions can often be achieved by reduced consideration to some subsystem of (X, Σ) if the relational homomorphism is only partially defined.)

[6] If the relation is algebraic and fully defined we say we have a "complete relational correspondence," which may or may not yield successful attempts at matched behavior. Algebraic techniques characterizing the construction of "covering homomorphisms" from such a correspondence yield methods useful in emulating behavior and computation (Nehaniv, 1996).

and elements $s_j \in S$ (for j from 1 to k), with

$$(s_j, t'_j) \in \varphi_2.$$

Each t'_i represents the product of zero or more of t_j's in the order they occur. That is, each s_j is related under the correspondence to the corresponding t'_j. (For mathematical convenience, one also defines t'_0 to be 1, the identity element of D, and s_0 to be 1, the identity element of S. These identity elements have no effect or duration.) Letting x_0 denote the starting state of system (X, Σ), $x_{j+1} = x_j \cdot s_{j+1}$ and $y_0 = y$, $y_k = y'$, and $y_{j+1} = y_j \cdot t'_{j+1}$, we require that

$$(x_j, y_j) \in \varphi_1 \quad \text{for all } 0 \leq j \leq k,$$

that is, the states attained by corresponding action-events must correspond according to the relational homomorphism φ.

We have that

$$x_j \xrightarrow{s_{j+1}} x_{j+1}$$

and

$$y_j \xrightarrow{t'_{j+1}} y_{j+1}$$

for all j $(0 \leq j < k)$. That is, formally, *the behaviors of the two systems correspond according to the relational morphism.*

Remark: The correspondence may be partial but enough of it must be defined to permit us to decide whether behaviors do or do not correspond. Notice that φ^{-1} must be defined on each of the y_j's $(0 \leq j \leq k)$, i.e., each y_j must be related to a state x_j of the imitator. Similarly, each factor action sequence t'_j must be related to a corresponding sequence s_j in the imitator.

A schematic view of the framework is shown in Figures 5 and 6. Evidence that bottle-nosed dolphins have solved a detailed correspondence problem between their own and human bodies has been demonstrated by Louis Herman at the Kewalo Marine Laboratory in Hawaii in experiments involving imitation of humans by dolphins (in press).

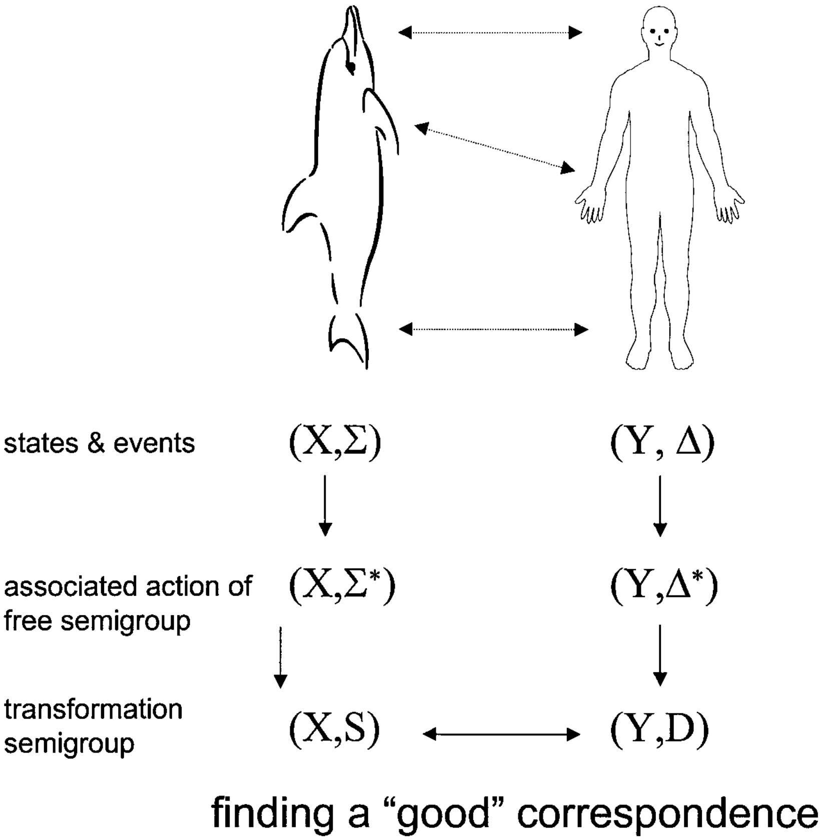

Figure 5. Schematic view on algebraic formulation of the correspondence problem: One applies the framework by generating automata models of the imitator and its model taking into account the body-environment coupling and events (actions and sensory input). From these two automata one derives two free transformation semigroups, whose inputs are sequences of events in the automata. Abstracting, one derives two transformation semigroups which abstract away particularities of state and action. The correspondence problem is a search for an appropriate relational homomorphism which encodes a correspondence between the imitator and model. Appropriateness of the correspondence is evaluated using metrics.

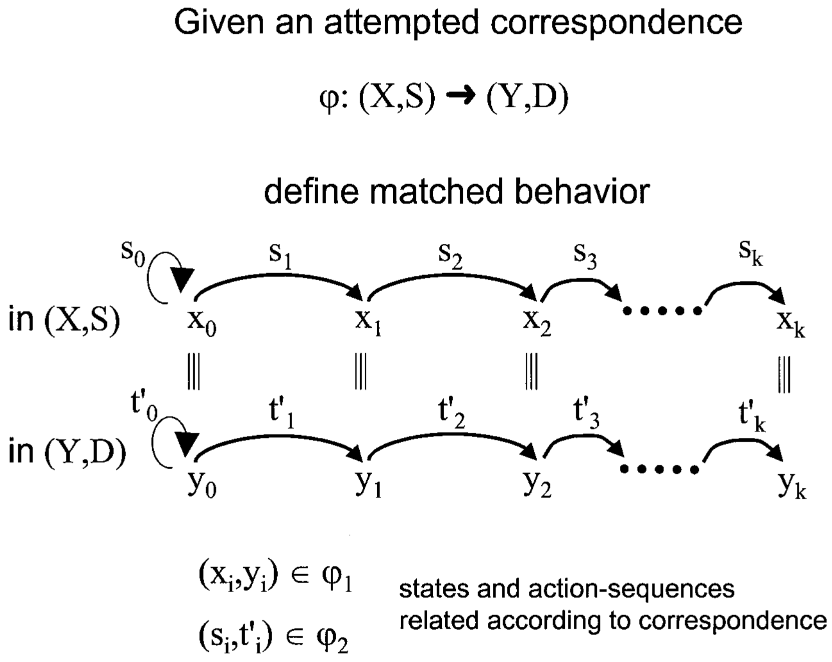

Figure 6. Schematic view on matched behaviors. A relational homomorphism can be used to guide attempts at matching behaviors. The horizontal transitions are corresponding action events in the two systems leading to corresponding states. The vertical links are correspondences as specified by the relational homomorphism. Finally (not shown), if the matched behavior is close enough for observer-attributed purposes (formalized using a metric *d*), then the attempted correspondence gives a successfully matched behavior.

But how 'good' is this correspondence of behaviors? This question is answered by a *metric*, which depends on the *effects* required of the matched behavior, and is addressed in the next section.

EVALUATING ATTEMPTED IMITATION USING METRICS

With the notation above, if each x_j is 'close' to y_j and each s_j and is 'close' to t'_j for the required purposes, then one has appropriate close correspondence of steps in both systems. This characterizes successful attempted matched behavior. But what is meant by 'close'?

Metrics

Now we describe one general method of quantifying the *error* or *degree of failure* of an attempted imitation for some purpose (see the section on dissimilar bodies, above). For simplicity, we shall assume here that the state sets X and Y and the event sequences Σ^* and Δ^* of the two systems are both contained in some larger state set Z of action-configurations.

A metric, as usually defined in mathematics, satisfies $d(x, y) = d(y, x)$, $d(x, y) \geq 0$, and $d(x, y) + d(y, z) \geq d(x, z)$. For example, distance in a plane is a metric, but so is the function, called the *discrete metric*, with $d(x, y) = 0$ if $x = y$ and $d(x, y) = 1$ otherwise. Unlike the usual definition *we allow distances to be zero for distinct pairs of points*. Thus these metrics are actually more "liberal" *pseudometrics* that need not distinguish between certain classes of states or event sequences. Throughout this article, "metric" should be understood as "pseudometric" in this sense.

Given a [pseudo]metric (distance measure) on state and action-event pairs

$$d : Z \times Z \to \mathrm{R}\,,$$

we can consider factorizations as in the previous section, which may or may not satisfy the condition for successful imitation. Then we have a measure of the "error" of the imitation defined, for a given factorization:

$$E_k = \sum_{i=0}^{k} d((x_i, s_i), (y_i, t_i')),$$

where k was the number of "granules" ((y_i, t_i')'s) in the factorization. Recall the y_i's here come from taking a factorization $t_1' \cdots t_k'$ of $t_1 \ldots t_m$ and defining y_0 to be the initial state of the model and $y_{i+1} = y_i \cdot t_{i+1}'$ for $i = 0, \ldots, k-1$. Computing the minimum of this quality over all factorizations yields a measure that is small for maximally successful attempts at imitation. E_k is zero for a successful imitation, for which the imitator's action state pairs ((x_i, s_i)'s) correspond to the k subgoals or effects characterized by ((y_i, t_i')'s) which were attained by model.

Metrics on States and on Event Sequences

Now the metric above evaluated the correspondence of state-event sequence pairs. If one is concerned with the matching of states (or results), then one needs to measure the correspondence in states. A metric on states

$$d_1 : X \times Y \rightarrow \mathrm{R}$$

does that, and no more.

If one is concerned with matching of actions, then one needs to measure the correspondence of action sequences. A metric on sequences

$$d_2 : \Sigma^* \times \Delta^* \rightarrow \mathrm{R}$$

does that, and no more.

Thus, a metric on state-event pairs can be obtained as a weighted sum of these two types of metrics:

$$d = w_1 d_1 + w_2 d_2,$$

where the weights $w_i \geq 0$ and $w_1 + w_2 = 1$.

So that we have a (pseudo)metric

$$d((x, \sigma), (y, \tau)) = w_1 d_1(x, y) + w_2 d_2(\sigma, \tau).$$

Types of Imitation: Weights and Granularity

A weight of w_1 set to zero means that metric gives no weight to states, whereas if w_2 is zero then the particular actions taken are given no weight. The former corresponds to goal or result emulation or stimulus enhancement, while the latter corresponds to action-level imitation, copying, or mimicry of actions. If granularities are restricted to $k = 1$ and $w_1 = 1$, then only the goal or resulting final state is important. As k increases, finer correspondence (subgoals for states or individual action events) becomes more important. By tuning the weights w_1 and w_2, one controls the relative importance of matching state results and matching actions. By tuning k one controls the coarseness or fineness of the matches (more subgoals and/or more actions). By choosing the metrics, one is choosing which states of the demonstrator are deemed to match those of the imitator and how closely they match (using d_1), and also which actions of the demonstrator match those of the imitator and degree of this matching (using d_2).

The closest imitation would attain an error value of 0 even when k is equal to m, the maximum level of granularity for m subgoals. Summing the E_k's gives a measure of error $E = \sum E_k$ at all levels of granularity. It is equal to zero for identical bodies performing identical actions in the same environment. It is also zero for successful correspondence, and is always nonzero for an attempted matched behavior that is not successful at some level of granularity. One could select between several candidate correspondences by requiring E_k or E be minimized. By choosing thresholds, $\varepsilon > 0$, and allowing $E < \varepsilon$ or $E_k < \varepsilon$, one could set the minimal acceptable level for successfully matched behaviors.

Then, a successful imitation for purposes of G is one for which the value of the error is zero for an appropriate choice of a sequence of subgoals—i.e. of effects of the actions and results of actions for an appropriate factorization. In contrast one attempted imitation as compared to another is worse, if its error is higher no matter what sequence of subgoals is chosen. Optimizing with respect to this measure can lead to attempted imitations that succeed or approximately succeed for purposes of G.

Such measures could be combined with evolutionary or computational optimization techniques in searches for relational homomorphisms that generate successful behavioral matches of the desired type.

Example: Discrete Metrics with Respect to a Purpose

Let $\gamma \subset X \times Y$ be the relation with $(x, y) \in \gamma$ if x and y are equivalent for purposes of G and, conversely, $(x, y) \notin \gamma$ if they are not. Similarly, let $\delta \subset \Sigma^* \times \Delta^*$ be the relation with $(\sigma, \tau) \in \delta$ if and only if σ is equivalent to τ for purposes of G. The attempted imitation is *successful for purposes of G* if, additionally, these corresponding states and events are equivalent for purposes of G:

$$(x_j, y_j) \in \gamma \quad \text{and} \quad (s_j, t'_j) \in \delta \quad \text{for all } 0 \leq j \leq k.$$

This is the simplest case of a single, unchanging metric and may be generalized in the obvious ways by varying γ and δ, with weighted sums depending on the stage in the behavior, or by using continuously varying metrics, etc.

For evaluating imitative behavior for purposes of G (see above), one may choose the metric $d = d_1 + d_2$ with $d_1(x, y) = 0$ if x and y correspond for purposes of G and $d_1(x, y) = 1$ if otherwise, and similarly with $d_2(s, t) = 0$ if s and t correspond for purposes of G and $d_2(s, t) = 1$ otherwise.

FURTHER APPLICATIONS AND VIEWPOINT ON CORRESPONDENCES

We close with a brief consideration of the relation between perception-action mappings and correspondences, and notions of imitation and behavior-matching in the larger scope of transferring useful information between distinct domains.

Perception-Action as an Instance of the Correspondence Problem

The problem of correspondence between systems and generating matching behaviors is a problem of finding a 'good' relational homomorphism (and if possible a covering homomorphism). We have focused on relating the two systems. In addition to directly relating these two systems, one could also (1) relate the imitator to its *perceptions* of the model (i.e., states and events in the model as perceived by the imitator), and also (2) relate the imitator to its own perceptions of itself acting in the world.

In regard to the second possibility, reinforcement and associative learning could then be employed in iterative construction of a relational homomorphism, and this technique could then also be applied to an agent-learning control of its body through perceptions of its actions in the environment. Thus formulated, the learning of motor control can be conceived of as the problem of imitating the self, whose actions and the effects of these actions are observed and used to generate a relational homomorphism which can be applied for purposeful behavior (i.e., to attain certain states and action-event effects—goals—in the body-environment coupling). This realization that sensory-motor control involves this kind of correspondence is present, for instance, in Heyes and Ray's associative sequence learning theory (ASL) in which the correspondence between visual stimuli and motor processes is encoded in "vertical links," while the successive events of an imitated behavior are given by “horizontal links” (Heyes & Ray, 2000). In these terms, the vertical links of ASL comprise a relational homomorphism, while the horizontal links encode attempted matched behaviors. Associative sequence learning theory has been used to approach the problem of *perceptual opacity*—i.e., how is it that one can imitate motor acts, such as bowing or tongue protrusion, for which sensory experience when observing them performed by others differs radically from feedback one experiences when performing them. Associative sequence learning theory, in its current form, does not address the effects of actions, but could be extended in line with ideas of effect-level imitation that we have discussed.

Metaphor and Imitation: Cross-Domain Transfer of Useful Information

Behavioral matching and imitation can be viewed as examples of *metaphor*, where metaphor is construed as mapping or synthesis of meaningful information between domains. This view of metaphor does not require it to have a linguistic or even conceptual substrate. Such a view of metaphor beyond language and concepts is developed in Nehaniv (1999a). Metaphor supports the transfer of meaningful or useful information from one domain to another. In this case of matched behavior, imitation, and observational learning, correspondence plays the role of the mapping, and metrics serve to evaluate attempted matched behavior—i.e., how good the metaphor between differently embodied systems is for purposes encoded in the metric.

ACKNOWLEDGMENTS

The work reported in this paper builds on work presented at the *European Workshop on Learning Robots 1998* (*EWLR*-*7*) in Edinburgh, Scotland on 20 July 1998 and in *Agents in Interaction—Acquiring Competencies Through Imitation Workshop* at *Autonomonous Agents ’98* in Minneapolis, MN in May 1998. It extends a portion of University of Aizu Technical Report 98-1-003 “Algebraic Engineering and Social Robotics” (27 April 1998) by the authors. A preliminary, less-refined version of this article is in *Learning Robots: An Interdisciplinary Approach*, J. Demiris and A. Birk, eds., World Scientific Press, River Edge, NJ.

This is Communication No. 1 of the University of Hertfordshire Adaptive Systems Research Group.

REFERENCES

Billard, A. 1999. *DRAMA, a connectionist model for robot learning: Experiments on grounding communication through imitation in autonomous robots.* Ph.D. thesis, Department of Artificial Intelligence, University of Edinburgh.

Billard, A., and K. Dautenhahn. 1997. Grounding communication in situated, social robots. In *Proc. "Towards Intelligent Mobile Robots" TIMR UK 97*. Technical Report Series of the Department of Computer Science, Manchester University, Manchester, UK.

Billard, A., and K. Dautenhahn. 1998. Grounding communication in autonomous robots: An experimental study. *Robotics and Autonomous Systems, Special Issue on "Scientific Methods in Mobile Robotics"* 24(1–2):71–81.

Billard, A., and G. Hayes. 1997. Learning to communicate through imitation in autonomous robots. In *Proceedings of Seventh International Conferences on Artificial Neural Networks* (*ICANN97*), eds. W. Gerstner et al., 763–768. New York: Springer-Verlag.

Breazeal, C. 1999. Imitation as social exchange between humans and robots. In *Proceedings of the AISB'99 Symposium on Imitation in Animals and Artifacts*, eds. K. Dautenhahn and C. L. Nehaniv, 96–104. Society for the Study of Artificial Intelligence and Simulation of Behaviour, Edinburgh, Scotland.

Breazeal, C. 2000. Sociable machines: Expressive social exchange between humans and robots. Sc.D. dissertation, Engineering and Computer Science, MIT, Cambridge, MA.

Breazeal, C., and B. Scassellati. Issues in building robots that imitate people. In *Imitation in animals and artifacts*, eds. K. Dautenhahn and C. L. Nehaniv. Cambridge, MA: MIT Press, (in press).

Brooks, R. 1996. Behavior-based humanoid robotics. In *Proc. 1996 IEEE/RSJ International Conference on Intelligent Robots and Systems, IROS 96*, 1–8.

Brooks, R. A., C. Breazeal, M. Marjanovic, B. Scassellati, and M. M. Williamson. 1999. The Cog project: Building a humanoid robot. In *Computation for metaphors, analogy, and agents*, ed. C. L. Nehaniv, 53–87. Springer LNAI, 1562.

Brooks, R. A., and L. A. Stein. 1994. Building brains for bodies. *Autonomous Robots* 1:7–25.

Byrne, R. W. 1999. Imitation without intentionality: Using string parsing to copy the organization of behavior. *Anim. Cogn.* 2:63–72.

Byrne, R. W., and A. E. Russon. 1997. Learning by imitation: A hierarchical approach. *Brain and Behavioral Sciences.* 21:667–709.

Call, J., and M. Carpenter. Three sources of social learning. In *Imitation in animals and artifacts*, eds. K. Dautenhahn and C. L. Nehaniv, Cambridge, MA: MIT Press, (in press).

Crutchfield, J. P. 1994. Observing complexity and the complexity of observation. In *Inside versus outside*, H. Atmanspacher, ed., 234–272. New York: Springer-Verlag.

Cypher, A. 1993. *Watch what I do: Programming by demonstration.* Cambridge, MA: MIT Press.

Dautenhahn, K. 1994. Trying to imitate—a step towards releasing robots from social isolation. In *Proc. From Perception to Action Conference, Lausanne, Switzerland*, September 7–9, eds. P. Gaussier and J.-D. Nicould. 290–301. Piscataway, NJ: IEEE Computer Society Press.

Dautenhahn, K. 1995. Getting to know each other—artificial social intelligence for autonomous robots. *Robotics and Autonomous Systems* 16:333–356.

Dautenhahn, K. 1997a. Biologically inspired robotic experiments on interaction and dynamic agent-environment couplings. In *Proc. Workshop SOAVE'97, Selbstorganization von Adaptivem Verhalten, Ilmenau*, September, 14–24.

Dautenhahn, K. 1997b. I could be you—the phenomenological dimension of social understanding. *Cybernetics and Systems* 25(8):417–453.

Dautenhahn, K. 1999. Embodiment and interaction in socially intelligent life-like agents. In *Computation for metaphors, analogy and agents*, C. L. Nehaniv, ed., 105–147. Springer LNAI, 1562.

Dautenhahn, K. and I. Werry. Issues of robot-human interaction dynamics in the rehabilitation of children with autism. In *Proc. From Animals to Animats: The Sixth International Conference on the Simulation of Adaptive Behavior* (*SAB2000*)*, 11–15 September*, Paris, France, J. A. Meyer, A. Berthoz, D. Floreano, H. Roitblat and S. W. Wilson, eds., 519–528. Cambridge, MA: MIT Press.

Demiris, J., and G. Hayes. 1997. Do robots ape? In *Socially Intelligent Agents: Papers of the AAAI Fall Symposium*, Cambridge, MA, November, 28–30. American Association for Artificial Intelligence.

Eilenberg, S. 1976. *Automata, languages and machines*, B. New York: Academic Press.

Gaussier, P., S. Moga, J. P. Banquet, and M. Quoy. 1998. From perception-action loops to imitation processes: A bottom-up approach of learning by imitation. *Applied Artificial Intelligence: An International Journal, Special Issue on "Socially Intelligent Agents,"* 12(7–8):701–729.

Gibson, J. J. 1977. *The ecological approach to visual perception.* Boston: Houghton Mifflin.

Hayes, G., and J. Demiris. 1994. A robot controller using learning by imitation. In *Proc. 2nd International Symposium on Intelligent Robotic Systems*, A.

Borkowski and J. L. Crowley, eds., 198–204. Grenoble, France, LIFTA-IMAG.

Herman, L. M. Vocal, social, and self imitation by bottlenosed dolphins. In *Imitation in animals and artifacts*, K. Dautenhahn and C. L. Nehaniv, eds. Cambridge, MA: MIT Press, (in press).

Heyes, C. M., and B. G. Galef Jr., eds. 1996. *Social learning in animals: The roots of culture.* New York: Academic Press.

Heyes, C. M., and E. Ray. 2000. What is the significance of imitation in animals? *Advances in the Study of Behaviour* 29:215–245.

Hinde, R. A., and J. Fisher. 1951. Further observations on the opening of milk bottles by birds. *Brit. Birds* 44:393–396.

Jones, J. L., and A. M. Flynn. 1993. *Mobile robots: Inspiration to implementation.* Natick, MA: A. K. Peters Ltd.

Lieberman, H. ed. 2000. Special section on ''Programming by Example.'' *Communications of the Association for Computing Machinery* 43(3):72–114.

McNeill Alexander, R. 1992. *Exploring biomechanics: Animals in motion.* San Francisco, CA: W.H. Freeman and Company.

Mitchell, R. W. 1987. A comparative-developmental approach to understanding imitation. *Perspectives in Ethology* 7:183–215.

Nehaniv, C. L. 1996. From relation to emulation: The covering lemma for transformation semigroups. *Journal of Pure and Applied Algebra* 107:75–87.

Nehaniv, C. L. ed. 1999a. *Computation for metaphors, analogy, and agents.* Springer LNAI, 1562.

Nehaniv, C. L. 1999b. Meaning for observers and agents. In *Proceedings of the 1999 IEEE International Symposium on Intelligent Control: Intelligent Systems & Semiotics*, 435–440. Institute of Electrical and Electronics Engineers, Cambridge, MA.

Nehaniv, C., and K. Dautenhahn. 1998a. Embodiment and memories—algebras of time and history for autobiographic agents. In *Cybernetics and Systems Research 98, Proc. 14th European Workshop on Cybernetics and Systems Research*—14–17 April, Vienna, Austria, Volume 2, 651–656. Austrian Society for Cybernetics Studies.

Nehaniv, C., and K. Dautenhahn. 1998b. Semigroup expansions for autobiographic agents. in *Proceedings of the First Symposium on Algebra, Languages and Computation*, University of Aizu, Japan, 30 October–1 November, 77–84. Osaka University.

Norman, D. A. 1992. *The psychology of everyday things.* London: Harper Collins.

Sammut, C., S. Hurst, D. Kedzier, and D. Michie. 1992. Learning to fly. In *Proc. Ninth International Conference on Machine Learning*, Aberdeen, Scotland, 385–393, Morgan Kaufman.

Thorndike, E. L. 1898. Animal intelligence: An experimental study of the associative process in animals. *Psychol. Rev. Monogr.* 2:551–553.

Tomasello, M. 1990. Cultural transmission in the tool use and communicatory signalling of chimpanzees? In *"Language" and intelligence in monkeys and apes: Comparative developmental perspectives*, S. Parker and K. Gibson. London: Cambridge University Press.

Tomasello, M., and J. Call. 1997. *Primate cognition.* Oxford University Press.

Visalberghi, E., and D. Fragaszy. "Do monkeys ape?" Ten years after. In *Imitation in animals and artifacts*, eds. K. Dautenhahn and C. L. Nehaniv. Cambridge, MA: MIT Press, (in press).

Whiten, A. Imitation of sequential and hierarchical structure in action: Experimental studies with children and chimpanzees. In *Proceedings of the AISB'99 Symposium on Imitation in Animals and Artifacts*, eds. K. Dautenhahn and C. L. Nehaniv, 38–46. Society for the Study of Artificial Intelligence and Simulation of Behaviour. Also in: *Imitation in animals and artifacts*, eds. K. Dautenhahn and C. L. Nehaniv. Cambridge, MA: MIT Press, (in press).

Zentall, T. R. 1996. An analysis of imitative learning in animals. In *Social learning in animals: The roots of culture*, eds. C. M. Heyes and B. G. Galef Jr., 221–243. New York: Academic Press.

Zentall, T. R. 2001. Imitation and other forms of social learning in animals. Evidence, function, and mechanisms. *Cybernetics and Systems*, Special Issue on Imitation in Natural and Artificial Systems. In this issue.